

## Обобщенные критерии идентификации химических соединений методами хроматографии – масс-спектрометрии

**\*Б.Л. Мильман<sup>1</sup>, И.К. Журкович<sup>2</sup>**

<sup>1</sup>Институт экспериментальной медицины,  
Российская Федерация, 197376, Санкт-Петербург, ул. Акад. Павлова, 12

<sup>2</sup>Институт токсикологии ФМБА России,  
Российская Федерация, 192019, Санкт-Петербург, ул. Бехтерева, 1

\*Адрес для переписки: Мильман Борис Львович, E-mail: [bmilman@mail.rcom.ru](mailto:bmilman@mail.rcom.ru)

Поступила в редакцию 1 июня 2020 г., после доработки – 17 июня 2020 г.

Представлен обзор критериев химической идентификации, применяемых в рамках целевого и нецелевого анализа методами хроматографии – масс-спектрометрии. Окончательные критерии в обоих случаях сопоставимы и предполагают близкое сходство характеристик хроматографического удерживания (времена или относительные времена) и масс-спектров (массы отдельных ионов, относительные интенсивности их пиков) идентифицируемых аналитов и аналитических стандартов. Сходство, как правило, означает максимально допустимые отклонения значений указанных величин, установленных для аналитов, от соответствующих справочных значений, приписанных известным соединениям. Используемые критерии идентификации включены в общие руководства по химическому анализу, подготовленные международными или национальными регулирующими организациями. В случае нецелевой идентификации аналиты и соответствующие стандарты неизвестны заранее. На первых этапах решения задач этого вида необходимо отобрать соединения – кандидаты на идентификацию, затем сравнить характеристики аналитов и соответствующих аналитических стандартов, используя критерии идентификации, аналогичные тем, которые сформулированы для целевого анализа. Для поиска кандидатов используют хроматографию и tandemную масс-спектрометрию высокого разрешения, продвинутое компьютерное моделирование и предсказания данных, поиски в химических базах данных, сводках индексов удерживания и масс-спектральных библиотеках. Применение этих методов и процедур – условие надежности поиска кандидатов на идентификацию. Прогресс в области нецелевого анализа зависит от усовершенствования масс-спектрометрической техники, в первую очередь за счет повышения точности и воспроизводимости определения масс ионов. Не меньшее значение имеет развитие компьютерных технологий (информатики), что должно обеспечить появление более полных библиотек масс-спектров и более точных методов предсказания масс-спектрометрических и хроматографических данных.

**Ключевые слова:** химическая идентификация, масс-спектрометрия, хроматография, компьютеры, количественные критерии идентификации.

For citation: *Analitika i kontrol'* [Analytics and Control], 2020, vol. 24, no. 3, pp. 164-173

DOI: 10.15826/analitika.2020.24.3.003

## Summarized criteria of chemical compounds identification by chromatography-mass spectrometry

**\*B.L. Milman<sup>1</sup>, I.K. Zhurkovich<sup>2</sup>**

<sup>1</sup> Institute of Experimental Medicine,  
ul. Akad. Pavlova 12, Saint Petersburg 197376, Russian Federation

<sup>2</sup> Institute of Toxicology,  
ul. Bekhtereva 1, Saint Petersburg 192019, Russian Federation

\*Corresponding author: Boris L. Milman, E-mail: [bmilman@mail.rcom.ru](mailto:bmilman@mail.rcom.ru)

Submitted 01 June 2020, received in revised form 17 June 2020

Criteria of chemical identification used in target and non-target analysis performed by chromatography-mass spectrometry are reviewed. In both cases, the final criteria are comparable and have a close similarity of both chromatographic retention values (times or relative times) and mass spectra (individual ion masses, their relative peak intensities) of identified analytes and analytical standards. Similarity, usually, means maximum permissible deviations of those values from the corresponding standard values assigned to the known compounds. Identification criteria used are contained in the general guides on the chemical analysis issued by the international or national regulatory organizations. In the case of non-target identification, the analytes and the corresponding standards are unknown before the analysis. In the first stages of this kind of analysis, the candidate compounds are selected for the identification and then the candidate features are compared to those of the standard compounds with the use of the criteria suggested for the target analysis. During the selection of the candidates, one uses high resolution chromatography – tandem mass spectrometry, advanced computer methods of data processing and prediction, and performs searches in chemical data bases, collections of retention indices, and mass spectral libraries. The use of these techniques and workflows can be considered as a reliable condition of the candidates' selection for the identification. The progress in non-target analysis depends on technique innovations in mass spectrometry, first, by improving the accuracy and reproducibility of ion masses. The development of computer technologies is no less important, which should ensure the emergence of more complete libraries of mass spectra and more accurate methods for predicting the mass spectrometric and chromatographic data.

**Keywords:** chemical identification, mass spectrometry, chromatography, computers, quantitative criteria of identification

## ВВЕДЕНИЕ

Химики, специализирующиеся в области молекулярного анализа, как правило, выбирают методы капиллярной газовой (ГХ) или высокоэффективной жидкостной (ВЭЖХ) хроматографии и масс-спектрометрии (МС) и их объединенные варианты (ХМС), что обусловлено высокой чувствительностью и селективностью этих методов. Их современное развитие связано с выпуском приборов новых типов, появлением новых компьютерных программ, общим прогрессом аналитической методологии и информатики [1]. Это же относится и к отдельным видам и областям химического анализа, таким как идентификация органических (биоорганических) соединений.

Индивидуальные химические соединения считают идентифицированными, если выполняются определенные критерии, прежде всего соответствие экспериментальных данных, полученных для аналитов, характеристикам (свойствам, значениям измеренных величин) тех или иных известных химических соединений [2,3]. В случае хроматомасс-спектрометрии к таким характеристикам/величинам относят времена или индексы удерживания (ИУ), массы ионов и относительные интенсивности их пиков, масс-спектры в целом. Критерии, как правило, связаны с допустимыми отклонениями значений указанных величин, установленных для аналитов, от соответствующих справочных значений, приписанных известным соединениям, а также с определенными требованиями к условиям хроматомасс-спектрометрического анализа. Допустимые отклонения зависят не только от свойств идентифицируемых соединений, но и разновидностей применяемых методов ХМС и вариантов аналитических задач.

Представления о критериях идентификации эволюционируют вместе с развитием аналитики. Наиболее полно, на взгляд авторов, требования к процедурам химической идентификации и ее критериям были изложены приблизительно 10 лет

тому назад в монографиях [2, 3]. Дальнейший прогресс химической аналитики, если иметь в виду ее идентификационные аспекты, удачно отражен в ряде новых руководств и обзоров, например в работах [4-9]. Тем не менее, специальные обобщающие работы в отношении модернизации критериев химической идентификации, насколько нам известно, в последние годы не появлялись.

Данная работа восполняет указанный пробел в обзорной литературе. Проблема критериев изложена в связи с устоявшимися представлениями о химической идентификации, основными вариантами ХМС, разновидностями аналитических процедур и общим прогрессом аналитики. Рассмотрены методы и критерии идентификации для двух основных видов анализа – целевого и нецелевого. Наиболее современные подходы к проведению качественного анализа, обуславливающие его качество, можно отнести к правилам «хорошей идентификационной практики» (ХИП, good identification practice [3]).

Критерии идентификации рассмотрены в *обобщенном* виде. Именно в таком виде они отражают результаты разработок методов и методик целевого анализа в отдельных лабораториях. Эти результаты *обобщены* в руководствах по проведению анализа в тех или иных областях (см. ниже). В случае нецелевого анализа, на первых его стадиях, критерии идентификации заменяются правилами (условиями) отбора перспективных версий (кандидатов на идентификацию). Такие правила, по существу, представляют собой необходимость применения наиболее современных *общих* аналитических методов и информационных процедур. Частные примеры использования различных критериев идентификации приведены в цитируемой литературе.

Объекты анализа в данной обзорной работе ограничены низкомолекулярными соединениями (типичные молекулярные массы – несколько сотен Да, вплоть до 1000-2000 Да).

## ХИМИЧЕСКАЯ ИДЕНТИФИКАЦИЯ. ПАРАДИГМА-2010

Особенности, подходы к разработке и реализации процедур химической идентификации, оценки ее правильности, сформулированные в первом десятилетии 21-го века [2, 3], представляют собой следующее.

1. По определению [2], «*идентификация (химическая) – отождествление аналита с известным химическим веществом (индивидуальным соединением, группой соединений), отнесение аналитического сигнала к известному веществу (соединению, группе соединений)*». Идентификация и качественный анализ – родственные понятия. Идентификация реализуется в соответствии с ее видами, способами, подходами, методами и методиками.

2. Существует несколько общих способов идентификации. (а) Наиболее надежный из них – совместный (или последовательный) анализ проб и соответствующих аналитических стандартов (стандартных образцов состава, веществ сравнения) и совпадение характеристик (значений измеряемых величин) идентифицируемых соединений и характеристик стандартов. В качестве последних часто используют (б) справочные данные экспериментального происхождения, полученные в сходных условиях, или (в) природные константы (формульные массы как суммы масс атомов, образующих молекулы, распространенности изотопов разных элементов). Менее надежны результаты, полученные с (г) расчетными (теоретическими) значениями в качестве справочных данных или (д) при интерпретации результатов аналитического эксперимента (спектров). Итак, надежно идентифицированные аналиты в своем большинстве – следствие применения способа (а). Во многих остальных случаях указывают на условную идентификацию; говорят о кандидатах на идентификацию, требующих подтверждения.

3. Наиболее надежные методы идентификации – методы ХМС. Они обеспечивают разделение малых количеств аналитов, отдельное их детектирование, их идентификацию за счет характеристичных масс-спектров, а также параметров хроматографического удерживания. МС широко применяется, что обусловлено распространением новых методов ионизации (прежде всего ионизации электрораспылением, **ИЭР**) и новых приборов в области tandemной масс-спектрометрии (**МС<sup>2</sup>**, **МС<sup>n</sup>**) и масс-спектрометрии высокого разрешения (**МСВР**), в т.ч. объединенных приборов (**МС<sup>2</sup>ВР**).

4. Задачи идентификации распадаются на два разных вида: целевой и нецелевой анализ. В первом случае аналиты известны заранее, и их определяют по методикам анализа (методикам выполнения измерений) в соответствии с заложенными

в них критериями идентификации. В случае ХМС критериями являются попадания времен ( $RT$ ) или относительных времен ( $RRT$ ) удерживания, масс ионов ( $m$ ) и относительных интенсивностей их пиков ( $I$ ) – в соответствующие допустимые интервалы этих значений для стандартов (соответственно  $\Delta RT$ ,  $\Delta RRT$ ,  $\Delta m$ ,  $\Delta I$ ). Рекомендуемые интервалы указанных переменных как критерии идентификации (критерии подтверждения идентичности) часто содержатся в общих аналитических руководствах международных или национальных регулирующих организаций.

5. Общепринятыми величинами, характеризующими надежность (правильность) применяемого метода идентификации определенного аналита, являются показатели правильных результатов, такие как показатели правильных положительных результатов (ПП):

$$ПП = \frac{100 \cdot n_{пп}}{n_{пп} + n_{ло}} \quad (1)$$

и показатель правильных отрицательных результатов (ПО):

$$ПО = \frac{100 \cdot n_{по}}{n_{по} + n_{лп}}, \quad (2)$$

где  $n_{пп}$ ,  $n_{по}$ ,  $n_{лп}$ ,  $n_{ло}$  – число соответствующих правильных и ложных результатов соответственно при многократном ( $N$  раз,  $N = n_{пп} + n_{по} + n_{лп} + n_{ло}$ ) повторении данного аналитического эксперимента. В биохимическом анализе (клинической диагностике) показатели ПП и ПО называют чувствительностью и специфичностью соответственно. Показатели ПП и ПО существенно ниже 100 % для неудовлетворительных методов или малых количеств аналитов (слабых аналитических сигналов).

6. Если состав анализируемой пробы, ее компоненты неизвестны в лаборатории, то соответствующий анализ называют нецелевым (другие термины: анализ пробы/образца неизвестного состава; unknown, non-target(-ed) или untargeted analysis). Методы нецелевого анализа основаны на применении ХМС и соответствующих систем справочной информации (библиотеки масс-спектров, сводки индексов удерживания). Показатели (1) и (2) пригодны для оценки эффективности таких методов; переменные  $n$  в этом случае означают долю из  $N$  тестовых соединений, идентифицированных правильно ( $n_{пп}$ ) или неправильно ( $n_{лп}$ ) при валидации/проверке соответствующего метода. Вероятность попадания разных аналитов (распространенных и редких соединений) в пробы, в общем случае, неодинакова. Байесовская статистика, учитывающая результаты предыдущих анализов или распространенность соединений, выраженную их цитируемостью в химической литературе, принимает во внимание указанную вероятность. При прочих равных условиях – сходстве хроматографических и спектральных характеристик нескольких потенциальных аналитов – химик может предпочесть более распространенные соединения.

## НОВЕЙШИЕ ТРЕНДЫ РАЗВИТИЯ АНАЛИТИКИ

Можно выделить несколько тенденций развития аналитики последнего десятилетия, не затрагивающих значительно базисные принципы химической идентификации, но увеличивающих эффективность соответствующих процедур и делающих их более распространенными.

1. Во-первых, следует отметить **развитие техники МС**. Все большее распространение получили tandemные масс-спектрометры высокого разрешения – сочетания квадрупольных и времяпролетных масс-анализаторов и появившиеся в середине первого десятилетия 21-го века приборы на основе орбитальной ионной ловушки [10]. Эти аналитические инструменты обеспечивают точность измерения масс ионов, составляющую малые единицы ppm и даже их доли, что привело к новому варианту термина MCBP: «масс-спектрометрия высокого разрешения с определением точной массы» (high resolution accurate mass mass spectrometry).

2. Отмеченному тренду сопутствует **развитие хроматографии**, направленное на увеличение эффективности разделения сложных смесей. Растет применение двумерной ГХ [11] и ультразэффективной жидкостной хроматографии (УЭЖХ, UPLC, HPLC) [12] в сочетании с MCBP. В случае полярных аналитов все чаще применяют HILIC-хроматографию [13].

3. **Развитие компьютерной техники**, технологии программирования, методов искусственного интеллекта, хемометрии и т.д. приобрело огромную важность. Оно обеспечило новые возможности обработки и распространения хроматомасс-спектрометрической информации, появление больших данных [1], прогнозирование хроматографических характеристик [14, 15] и масс-спектров (*in silico* масс-спектры) [14, 16], создание баз данных [1, 14, 17] и электронных библиотек масс-спектров [1, 14, 18].

4. Достигнут значительный **прогресс биомедицинского анализа**, который выражается в широком развитии геномики, протеомики и, главное, метаболомики [4, 5, 8, 13-14], относящейся к тематике обзора. Многие методические достижения в сфере метаболомики могут быть распространены на другие области анализа низкомолекулярных соединений (анализ объектов окружающей среды и природных объектов, судебная медицина, токсикология).

Развитие аналитики проявилось в отношении обоих видов анализа/идентификации.

## ЦЕЛЕВОЙ АНАЛИЗ

Рекомендации по проведению целевого определения аналитов изложены в общих руководствах и методиках анализа, разрабатываемых в соответствии с этими руководствами. Рекомендуемые критерии идентификации, обобщающие результаты работы профильных лабораторий и содержащиеся в ру-

ководствах последнего десятилетия, приведены в табл. 1 и 2. Принцип анализа показан на рис (слева).

В табл. 1 содержатся хроматографические критерии, представляющие собой, как правило, максимально допустимое отклонение времени

Таблица 1

Хроматографические критерии идентификации

Table 1

Chromatographic identification criteria

Организация, документ	Аналиты, матрицы	Критерий
Европейский союз (EU) [19]* 2019 г.	Остатки пестицидов в продуктах и кормах	$\Delta RT$ , $\pm 0.1$ мин
Всемирное антидопинговое агентство (WADA) [20] 2015 г.	Допинг в биопробах	$\Delta RT^*$ , $\pm$ полуширина пика, или $\pm 0.1$ мин или $\pm 1\%^{3*}$ ; или $\Delta RRT^*$ , $\pm 1\%$ или $\pm 0.5\%^{4*}$
Управление по контролю качества пищевых продуктов и медикаментов США (FDA) [21] <sup>5*</sup> 2015 г.	Остатки ветеринарных лекарств, пестицидов, красителей и др. в продуктах, кормах, пищевых добавках, косметике и др.	$\Delta RT$ , $\pm 0.2$ мин или $\pm 2.5\%$ (но $\leq 0.5$ мин) или в пределах погрешности $RT$ (но $\leq 0.5$ мин) <sup>6*</sup>
Ассоциация химиков в конном спорте (AORC) [22] 2016 г.	Допинг в биопробах	ГХ: $\Delta RT$ , $\pm 1\%$ или $6\text{ с}^{7*}$ или $\Delta RRT$ , $\pm 1\%$ ВЭЖХ: $\Delta RT$ , $\pm 1\%$ или $12\text{ с}^{7*}$ или $\Delta RRT$ , $\pm 2\%$ УЭЖХ: $\Delta RT$ , $\pm 50\%$ полуширины пика стандарта или $3\text{ с}^{7*}$

**Примечания:** \* – допускаются большие отклонения: для изотопно-меченых внутренних стандартов или в других случаях, обоснованных при валидации методик. Пики аналита на разных масс-хроматограммах аналита должны совпадать по времени; <sup>2\*</sup> – в сравнении с объединенным пиком аналита и его аналитического стандарта при смешении их растворов (spiked sample); <sup>3\*</sup> – наибольшая из этих двух величин и при условии, что они меньше полуширины пика аналита в смешанном растворе; <sup>4\*</sup> – второе значение в случае изотопно-меченых внутренних стандартов; <sup>5\*</sup> – пики аналита на разных масс-хроматограммах должны совпадать по времени. Следует учитывать матричные эффекты, меняющие времена удерживания; <sup>6\*</sup> – критерии устанавливаются при валидации методик; <sup>7\*</sup> – наибольшая из этих двух величин



удерживания ( $\Delta RT$ ) или относительного времени удерживания ( $\Delta RRT$ ) идентифицируемого аналита от этой величины для аналитического стандарта. Выполнение критериев устанавливается в параллельных экспериментах или при смешении анализируемой пробы и стандарта. По сравнению с критериями десятилетней давности [2, 3] обращает

на себя внимание то, что для газовой и жидкостной хроматографии в большинстве случаев предложены одинаковые критерии (табл. 1). Это обстоятельство связано с успехами УЭЖХ и доступностью приборов и соответствующих колонок, обеспечивающих узкие хроматографические пики, сопоставимые с сигналами в капиллярной ГХ.

Таблица 2

Масс-спектрометрические критерии идентификации

Table 2

Mass spectrometric identification criteria

Документ	(ИЭ)-МС <sup>1</sup> , полные сканы	(ИЭ)-МС <sup>1</sup> , SIM	(ИЭР)-МС <sup>2(n)</sup>	МСВР
EU [19]* 2019 г.	Справочный спектр регистрируется на том же приборе, лучше в той же серии анализов. Порог сходства спектров определяется в лаборатории	Три иона, желателен в т.ч. молекулярный ион <sup>2*</sup> , $\Delta I$ , $\pm 30$ %	SRM: два фрагментных иона, $\Delta I$ , $\pm 30$ % (отн.)	Два иона, желательны протонированная (или катионированная) молекула и ион-продукт, $\Delta m$ , $\pm 5$ ppm ( $< 1$ мДа для $m/z < 200$ ) МС <sup>2</sup> : оценки $\Delta I$ необязательны
WADA [20]*,3*,4* 2015 г.	Минимум три иона. $\Delta I$ : $\pm 10$ (абс.) при $I = 50-100$ %, $\pm 20$ % (отн.) при $I = 25-50$ %, $\pm 5$ (абс.) при $I = 1-25$ %		SRM: Минимум два фрагментных иона, $\Delta I$ , см. (ИЭ)-МС <sup>1</sup> ,	
FDA [21, 23]* 2015 г.	Минимум три иона, без слабых сигналов. «Общее соответствие» значений / аналита и стандарта	Три иона: $\Delta I$ , $\pm 10$ % (отн.). Четыре и более иона: $\Delta I$ , $\pm 15$ % (отн.). Изотопные пики или потеря $H_2O$ : $\Delta I$ , $\pm 10$ % (отн.)	Полные сканы: см. (ИЭ)-МС <sup>1</sup> , SRM: три и более иона, $\Delta I$ , $\pm 30$ % (отн.) <sup>5*</sup>	МС <sup>1</sup> , SIM: оптимизация окна выделения масс. МС <sup>1</sup> : два и более ионов, $\Delta m$ , $\pm 5$ ppm. МС <sup>2</sup> : два и более ионов, $\Delta m$ , $\pm 10$ ppm. МС <sup>1</sup> и МС <sup>2</sup> : два и более ионов, $\pm 5$ ppm и $\pm 10$ ppm соответственно. Дополнения <sup>6*</sup>
AOPC [22]*,4* 2016 г.	Минимум три иона, желателен в т.ч. молекулярный ион (если $I > 5$ %). $\Delta I$ , $\pm 10$ % (абс.) или $\pm 30$ % (отн.) <sup>7*</sup> . Присутствие в спектре всех ионов с $I > 10$ %	Минимум четыре иона: $\Delta I$ , $\pm 10$ % (абс.) или $\pm 25$ % (отн.) <sup>7*</sup>	Полные сканы: минимум три иона. $\Delta I$ , $\pm 20$ % (абс.) или $\pm 40$ % (отн.) <sup>7*</sup> . Присутствие в спектре пиков всех ионов с $I > 10$ %. При недостатке фрагментных ионов учитывается протонированная или катионированная молекула, если $I = 10-80$ %. SRM: минимум три фрагментных иона, $\Delta I$ , $\pm 10$ % (абс.) или $\pm 30$ % (отн.) <sup>7*</sup>	Аналогично требованиям низкого разрешения. При узких окнах масс ( $\pm 5$ ppm или $\pm 2$ мДа <sup>7*</sup> ) оценка $\Delta I$ не применяется

**Примечания:** \* —  $S/N$  не менее или более 3 : 1; во всех случаях имеются ввиду характеристичные ионы; <sup>2\*</sup> — для более уверенных умозаключений могут потребоваться другие ионы, в т.ч. изотопная картина; <sup>3\*</sup> — отклонение от соответствующих масс ионов аналитического стандарта  $\Delta m$  - не более  $\pm 0.5$  Да; <sup>4\*</sup> — применение другого метода ионизации/фрагментации или дериватизация, если не находится требуемое количество ионов; <sup>5\*</sup> — при полной фрагментации иона-предшественника в спектре МС<sup>n</sup> рекомендуются два дополнительных характеристичных иона, зарегистрированных в условиях МС<sup>n+1</sup>; <sup>6\*</sup> — оптимизация окна выделения масс при МС<sup>1</sup>, SIM; оценки  $\Delta I$  как критерия при выходе из границ  $\Delta m$ ; <sup>7\*</sup> — наибольшая из этих двух величин.

Табл. 2 включает масс-спектрометрические критерии. Они представляют собой требования, во-первых, к полным масс-спектрам (полным сканам)  $MS^1$  (летучие соединения, используется ионизация электронами, ИЭ) и  $MS^{2(n)}$ . Tandemная MS с сочетанием ионизации электрораспылением (ИЭР) и фрагментации, активируемой соударениями, как правило, используется в случае нелетучих аналитов. Во-вторых, эти требования относятся к режимам мониторинга выбранных ионов (SIM в  $MS^1$ ) и реакций (MRM в  $MS^{2(n)}$ ). Среди критериев: максимально допустимое отклонение масс ионов ( $\Delta m$ ) и относительных интенсивностей ( $\Delta I$ ) их пиков в случае идентифицируемого аналита – от этих величин для аналитического стандарта. По сравнению с требованиями к идентификации, характерными для аналитической практики предшествующего десятилетия, введены некоторые изменения и уточнения, прежде всего касающиеся MCBP (табл. 2). Так, учтено появление и широкое использование  $MS^2BP$ , причем становится необязательным критерий, связанный с  $\Delta I$ . Причина этого, по-видимому, заключается в недостаточной воспроизводимости этих величин; отметим, что при идентификации по справочным масс-спектрам высокого разрешения точные значения масс  $m$  гораздо важнее, чем интенсивности пиков  $I$  [24].

Однотипные критерии, внесенные в табл. 1 и особенно 2, не полностью тождественны, например, отличаются шириной интервалов значений переменных. Это отражает несовпадающий опыт работы различных лабораторий, на основе которого сформулированы требования к методам ХМС, и разную природу аналитов и матриц. Поэтому можно считать, что у аналитика есть некоторая свобода выбора критериев. Тем не менее, для идентификации обязательно соответствие как хроматографическим, так и масс-спектрометрическим критериям, зависящим от природы аналитов и применяемых методов ХМС. Необходимо обратить внимание и на уточнения условий идентификации (примечания к табл. 1 и 2).

При разработке методик целевого определения соединений других групп, не отмеченных в табл. 1, или адаптации уже разработанных методик следует верифицировать рассмотренные хроматографические и масс-спектрометрические критерии и валидировать их, если необходима модификация.

## НЕЦЕЛЕВОЙ АНАЛИЗ: ХОРОШАЯ ИДЕНТИФИКАЦИОННАЯ ПРАКТИКА

Ввиду огромного количества известных соединений (десятки млн. [2, 3, 17]) и разброса их свойств, создание подробных руководств по общему химическому анализу (нецелевой химической идентификации) выглядит недостижимой задачей. Тем не менее, целесообразно выделить и охарактеризовать общие стадии нецелевого анализа, которые в

своем большинстве, так или иначе, необходимы для надежного поиска и идентификации неизвестных компонентов анализируемых проб. Краткие описания этих стадий, представляющих собой компоненты ХИП, учитывают прогресс  $MS^2BP$  и начинают попадать в рассмотренные выше общие аналитические руководства (документ FDA [21]). В соответствии с руководством [21] выделяют следующие стадии:

- хроматографическое разделение,
- выделение масс-спектров отдельных аналитов, отдельных сигналов в этих спектрах (массовых пиков) и другие компьютерные процедуры,
- генерация молекулярных формул,
- поиск в базах данных.

Многочисленные обзоры, например [4-9], дают более полную информацию о процедурах нецелевого анализа, которые в коротком изложении, за исключением стадий пробоотбора и пробоподготовки (о них см. литературу в обзоре [17, Table 1]), выглядят следующим образом.

**Хроматографическое разделение.** Различные компоненты анализируемых образцов должны быть максимально разделены, чтобы корректно определить параметры удерживания и выделить масс-спектры отдельных соединений. В связи с этим возрастает применение новых методов хроматографии (см. выше) и других методов разделения в сочетании MCBP. Что касается параметров удерживания, в ГХ широко распространены справочные ИУ экспериментального происхождения [25].

**Компьютерные процедуры.** Работа с образцами сложного состава, какими предстают многие биомедицинские и пищевые пробы и объекты окружающей среды, требует компьютерной обработки получаемых хроматомасс-спектрометрических данных с использованием специальных программ, которыми снабжены современные приборы. Так, необходимы удаление фоновых сигналов и деконволюция хроматографических и массовых пиков – на отдельные компоненты, соответствующие индивидуальным аналитам. При этом реализуется специальная процедура – отбор/обнаружение/выделение массовых пиков (peak picking). При поиске подозреваемых (suspected) компонентов смесей может потребоваться генерация отдельных масс-хроматограмм, содержащих пики ионов, выбранные в узком окне масс (интервале значений  $m/z$ ) по всей длительности хроматограммы.

Учитывая многокомпонентный состав биопроб (многие тысячи и даже десятки тысяч соединений), аналитикам необходимы программы быстрой автоматической идентификации компонентов проб. Во многом такой анализ является скринингом [2, 3], идентификация носит условный (предварительный) характер, и здесь употребляется популярный в англоязычной литературе термин «аннотация» (аннотация аналитических сигналов).

**Генерация молекулярных формул.** Программы еще одной направленности подбирают сочетания

Таблица 3

Большие химические БД

Table 3

Big chemical databases

БД	Число соединений и веществ*, млн	Свободный доступ
CAS [26]	160	-
PubChem [27]	102.7 (соединения) 253.4 (соединения и вещества)	+
ChemSpider [28]	81	+
ZINC15** [29]	более 230 (оценка)	+

**Примечания:** \* – вещество - смесь индивидуальных соединений; \*\* – сводный электронный каталог реактивов и химикатов.

атомов (C, H, N, O и др.), соответствующие сходству точно измеренных молекулярных масс (выводятся из масс ионов  $M^+$ ,  $[M+H]^+$ ,  $[M+Na]^+$ ,  $[M-H]^+$  и др., обнаруженных в режиме МС<sup>1</sup>) и разных формульных масс. Разность таких экспериментальных и теоретических величин при ее малом значении (единицы или доли ppm) является показателем правильности определения молекулярной формулы. Если молекулярные массы достигают уровня сотен Да и более, то наблюдается много вариантов идентифицируемых формул. Предстоит их обширный перебор в поисках единственного правильного варианта и дальнейший выбор из различных изомеров, имеющих одну формулу.

**Поиски в химических базах данных.** По найденным молекулярным формулам (или даже массам ионов  $[M+H]^+$  и др.) проводят поиск в больших химических базах данных (БД) [1, 17, 30]. Наиболее

крупные БД этого вида указаны в табл. 3. Результаты поиска (индивидуальные аналиты) ранжируют по степени их популярности/распространенности, измеряемой соответствующей цитируемостью – числом баз данных или научных публикаций, содержащих информацию о конкретном соединении, другими словами, количеством информации о нем [30]. Наиболее распространенные соединения выделяют в качестве кандидатов на идентификацию. Их число можно сократить за счет веществ, не подходящих по свойствам, получению и применению.

#### Поиски в библиотеках масс-спектров.

Зарегистрированные фрагментные масс-спектры (приборы ИЭ-МС<sup>1</sup> и ИЭР-МС<sup>2(n)</sup>) сравнивают со справочными спектрами, помещенными в библиотеки масс-спектров (табл. 4). Соединения (или одно из них) с наиболее похожими спектрами, при отсутствии противоречий с данными МСВР и сведениями, содержащимися в химических БД (распространенность, физико-химические свойства, получение, применение), рассматриваются как кандидаты на идентификацию.

Выявленные соединения-кандидаты «пропускают» через фильтры справочных хроматографических данных (при условии, что они имеются). Если остается одно из соединений (другие резко отличаются по характеристикам), его можно считать условно идентифицированным. Этого может быть достаточно для решения многих аналитических задач. Если такие задачи носят особенно ответственный характер (сопутствуют катастрофам, массовым отравлениям или болезням и т.п.), или остаются несколько соединений-кандидатов, для однозначной идентификации требуются аналитические стандарты. Сравнивают хроматомасс-спектрометрические данные для аналитов и стандартов. Их совпадение

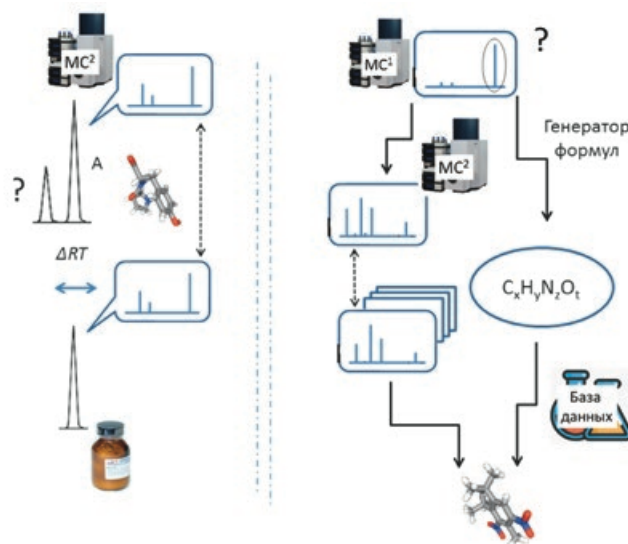
Таблица 4

Большие библиотеки масс-спектров

Table 4

Big mass spectral libraries

Соединения	Название	Спектры	Соединения	Комментарии
Летучие соединения, ЭИ-МС <sup>1</sup>	Wiley Registry 12 <sup>th</sup> [31]	817290	668435	Коммерческий продукт
	NIST 20 [32]	350643	306869	Коммерческий продукт
	NIST 20, МС <sup>2</sup> [32]	1,3 млн	31 тыс.	Коммерческий продукт
Нелетучие соединения, (ИЭР)-МС <sup>2(n)</sup>	METLIN [33]		> 300 тыс.	Преобладают метаболиты. Поиск на сайте [34]. Также присутствуют <i>in silico</i> масс-спектры
	The Global Natural Product Social Molecular Networking (GNPS) [35]	221083	18163	Природные соединения
	MassBank of North America (MoNA) [36]	124457		Преобладают спектры биологически активных соединений. Включены <i>in silico</i> масс-спектры
	MassBank [37]	60 тыс.	≤ 14 тыс	МС <sup>n</sup> , биологически активные соединения



**Рис.** Схема целевого (левая часть, аналит А) и нецелевого (справа, неизвестное соединение) анализа. Целевой анализ реализуется в соответствии с хроматографическим и масс-спектрометрическим критериями и наличием аналитического стандарта А. На хроматограмме присутствует второй значительный пик, принадлежность которого определенному соединению устанавливается по правилам нецелевого анализа (см. текст). Здесь показан упрощенный вариант идентификации: одно соединение с совпадающими результатами поисков в библиотеке спектров  $MS^2$  и химической базе данных для одной из молекул.

**Fig.** Scheme of target (left side, analyte A) and non-target (right side, unknown compound) analysis. Target analysis is carried out according to the chromatographic and mass spectrometric criteria and the availability of the analytical standard A. The second significant peak is present on the chromatogram, the identity of which is determined by the principles of non-targeted analysis (see text). A simple identification kind is shown here: one compound with matching searches in both the  $MS^2$  library and the chemical database for one molecule.

по критериям, рассмотренным выше или аналогичным им, определяет однозначную идентификацию аналита.

Решению данной задачи, однако, могут препятствовать некоторые обстоятельства, например, появление редкого, малоизученного аналита. При этом возможны проблемы с получением аналитического стандарта, и приходится привлекать другие методы для идентификации аналита [3]. В общем случае нельзя также исключать неизвестные ранее химические превращения аналитов на стадии пробоподготовки и проведения анализа, что влечет за собой получение ложных результатов идентификации. Они могут быть исключены при варьировании условий подготовки проб и самого анализа в ходе разработки или адаптации методики определения редких соединений. При этом проверяют и/или изменяют критерии идентификации (табл. 1 и 2).

Общая схема нецелевого анализа в упрощенном варианте показана на рисунке.

## ЗАКЛЮЧЕНИЕ

Задачи химической идентификации распадаются на два класса – целевой и нецелевой анализ. Первый вид обсуждаемого качественного анализа представляется относительно несложной проблемой при условии, что имеются необходимые приборы и реактивы, а персонал лаборатории обладает необ-

ходимой квалификацией. При этом используются обобщенные критерии идентификации, которые включены в разные руководства по химическому анализу.

Нецелевой анализ – гораздо более сложная задача, находится на грани науки и искусства (искусства проведения научных исследований). Для поиска кандидатов на идентификацию используются принципы ХИП, т.е. современная аналитическая методология, включающая применение хроматографии и масс-спектрометрии высокого разрешения, продвинутые методы компьютерной обработки и предсказания данных, поиски в химических базах данных и спектральных библиотеках. Обязательное применение этих методов и процедур представляет собой условие (по существу, критерий) надежного поиска соединений – кандидатов на идентификацию. Критерии же самой идентификации, сформулированные для целевого анализа, здесь применимы на последней стадии, когда сформирован список соединений-кандидатов, подобраны соответствующие аналитические стандарты и, в целом, уточнены и валидированы условия анализа.

## ЛИТЕРАТУРА

1. Мильман Б.Л., Журкович И.К. Большие данные в современном химическом анализе // Ж. аналит. химии. 2020. Т. 74, № 4. С. 316-326.



2. Мильман Б.Л. Введение в химическую идентификацию. СПб: ВВМ, 2008. 180 с.
3. Milman B.L. Chemical identification and its quality assurance. Berlin: Springer, 2011. 281 p.
4. Watson D.G. A rough guide to metabolite identification using high resolution liquid chromatography mass spectrometry in metabolomic profiling in metazoans // *Comput. Struct. Biotechnol. J.* 2013. V. 4, № 5. P. e201301005.
5. Proteomics, lipidomics, metabolomics: a mass spectrometry tutorial from a computer scientist's point of view / R. Smith [et al.] // *BMC Bioinf.* 2014. V. 15, № 7. P. S9.
6. Milman B.L. General principles of identification by mass spectrometry // *TrAC Trends Anal. Chem.* 2015. V. 69. P. 24-33.
7. A tutorial in small molecule identification via electrospray ionization-mass spectrometry: The practical art of structural elucidation / T. De Vijlder [et al.] // *Mass Spectrom. Rev.* 2018. V. 37, № 5. P. 607-629.
8. Metabolite Annotation and Identification / J. Godzien [et al.] // *Compr. Anal. Chem.* 2018. V. 82. P. 415-445.
9. Guidelines for unequivocal structural identification of compounds with biological activity of significance in food chemistry (IUPAC Technical Report) / R. J. Molyneux [et al.] // *Pure Appl. Chem.* 2019. V. 91, № 8. P. 1417-1437.
10. Zubarev R.A., Makarov A. Orbitrap mass spectrometry // *Anal. Chem.* 2013. V. 85, № 11. P. 5288-5296.
11. Comprehensive two-dimensional gas chromatography in forensic science: A critical review of recent trends / B. Gruber [et al.] // *TrAC Trends Anal. Chem.* 2018. V. 105. P. 292-301.
12. Nováková L., Svoboda P., Pavlík J. Ultra-high performance liquid chromatography // *Liquid Chromatography. Fundamentals and Instrumentation.* Elsevier, 2017. P. 719-769.
13. HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS / D.Q. Tang [et al.] // *Mass Spectrom. Rev.* 2016. V. 35, № 5. P. 574-600.
14. Software tools and approaches for compound identification of LC-MS/MS data in metabolomics [I. Blaženović et al.] // *Metabolites.* 2018. V. 8, № 2. P. C. 31.
15. Various aspects of retention index usage for GC-MS library search: A statistical investigation using a diverse data set / [Matyushin D.D. et al.] // *Chemom. Intell. Lab. Syst.* 2020. P. 104042.
16. Milman B.L., Ostrovidova E.V., Zhurkovich I.K. Isomer differentiation using *in silico* MS<sup>2</sup> spectra. A case study for the CFM-ID mass spectrum predictor // *Mass Spectrom. Lett.* 2019. V. 10, № 3. P. 93-101.
17. Milman B.L., Zhurkovich I.K. The chemical space for non-target analysis // *TrAC Trends Anal. Chem.* 2017. V. 97. P. 179-187.
18. Milman B. L., Zhurkovich I. K. Mass spectral libraries: A statistical review of the visible use // *TrAC Trends Anal. Chem.* 2016. V. 80. P. 636-640.
19. Analytical quality control and method validation procedures for pesticide residues analysis in food and feed. [Электронный ресурс]: [https://ec.europa.eu/food/sites/food/files/plant/docs/pesticides\\_mrl\\_guidelines\\_wrkdoc\\_2019-12682.pdf](https://ec.europa.eu/food/sites/food/files/plant/docs/pesticides_mrl_guidelines_wrkdoc_2019-12682.pdf). (дата обращения 10.05.2020).
20. Minimum criteria for chromatographic-mass spectrometric confirmation of the identity of analytes for doping control purposes. [Электронный ресурс]: [https://www.wada-ama.org/sites/default/files/resources/files/td2015idcr\\_-\\_eng.pdf](https://www.wada-ama.org/sites/default/files/resources/files/td2015idcr_-_eng.pdf). (дата обращения 10.05.2020).
21. Acceptance criteria for confirmation of identity of chemical residues using exact mass data for the FDA foods and veterinary medicine program. [Электронный ресурс]: <https://www.fda.gov/media/96499/download>. (дата обращения 10.05.2020).
22. AORC guidelines for the minimum criteria for identification by chromatography and mass spectrometry. [Электронный ресурс]: <http://www.aorc-online.org/documents/aorc-ms-criteria-modified-23-aug-16>. (дата обращения 10.05.2020).
23. Guidance for industry: Mass spectrometry for confirmation of the identity of animal drug residues. [Электронный ресурс]: <https://www.fda.gov/media/70154/download>. (дата обращения 10.05.2020).
24. Milman B.L., Zhurkovich I.K. Identification performance of low-molecular compounds by searching tandem mass spectral libraries with simple peak matching // *Mass Spectrom. Lett.* 2018. V. 9, № 3. P. 73-76.
25. NIST 17 GC method / Retention index library. [Электронный ресурс]: <https://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:ridatabase> (дата обращения 10.05.2020).
26. CAS content. [Электронный ресурс]: <https://www.cas.org/about/cas-content> (дата обращения 10.05.2020).
27. PubChem. [Электронный ресурс]: <https://pubchem.ncbi.nlm.nih.gov> (дата обращения 10.05.2020).
28. ChemSpider. [Электронный ресурс]: <http://www.chemspider.com> (дата обращения 10.05.2020).
29. ZINC 15. [Электронный ресурс]: <http://zinc15.docking.org> (дата обращения 10.05.2020).
30. Мильман Б.Л., Островидова Е.В., Журкович И.К. Большие химические базы данных свободного доступа в нецелевом масс-спектрометрическом анализе // *Масс-спектрометрия.* 2020. Т.17, № 2. С. 125-132.
31. Wiley registry of mass spectral data. [Электронный ресурс]: <https://www.sisweb.com/software/wiley-registry.htm> (дата обращения 10.05.2020).
32. The NIST 20 Mass spectral library. [Электронный ресурс]: <https://www.sisweb.com/software/ms/nist.htm> (дата обращения 10.05.2020).
33. Montenegro-Burke J.R., Guigas C., Siuzdak G. METLIN: a tandem mass spectral library of standards // *Computational Methods and Data Analysis for Metabolomics.* New York, Humana. 2020. P. 149-163.
34. METLIN. [Электронный ресурс]: [https://metlin.scripps.edu/ms\\_ms\\_spectrum\\_match\\_search.php](https://metlin.scripps.edu/ms_ms_spectrum_match_search.php) (дата обращения 10.05.2020).
35. GNPS: Global natural products social molecular networking. [Электронный ресурс]: <https://dorresteinlab.ucsd.edu/gnps> (дата обращения 10.05.2020).
36. MoNA - MassBank of North America. [Электронный ресурс]: <https://mona.fiehnlab.ucdavis.edu> (дата обращения 10.05.2020).
37. MassBank. [Электронный ресурс]: <http://www.massbank.jp> (дата обращения 10.05.2020).

## REFERENCES

1. Milman B.L., Zhurkovich I.K. Big data in modern chemical analysis. *J. Anal. Chem.*, 2020, vol. 75, no. 3, pp. 443-452. DOI: [org/10.1134/S1061934820020124](https://doi.org/10.1134/S1061934820020124).
2. Milman B.L. *Vvedenie v khimicheskuiu identifikatsiiu [Introduction to Chemical Identification]*. SPb: VVM, 2008. 180 p. (in Russian).
3. Milman B. L. *Chemical identification and its quality assurance*. Berlin: Springer, 2011. 281 p.
4. Watson D.G. A rough guide to metabolite identification using high resolution liquid chromatography mass spectrometry in metabolomic profiling in metazoans. *Comput. Struct. Biotechnol. J.*, 2013, vol. 4, no. 5, p. e201301005. Available at: <https://doi.org/10.5936/csbj.201301005> (Accessed 20 May, 2020).

5. Smith R., Mathis A.D., Ventura D., Prince J.T. Proteomics, lipidomics, metabolomics: a mass spectrometry tutorial from a computer scientist's point of view. *BMC Bioinf.*, 2014, vol. 15, no. 7, p. S9. Available at: <https://doi.org/10.1186/1471-2105-15-S7-S9> (Accessed 20 May, 2020).
6. Milman B.L. General principles of identification by mass spectrometry. *TrAC Trends Anal. Chem.*, 2015, vol. 69, pp. 24-33. DOI: [org/10.1016/j.trac.2014.12.009](https://doi.org/10.1016/j.trac.2014.12.009).
7. De Vijlder T., Valkenburg D., Lemi re F., Romijn E.P., Laukens K., Cuyckens F. A tutorial in small molecule identification via electrospray ionization - mass spectrometry: The practical art of structural elucidation. *Mass Spectrom. Rev.*, 2018, vol. 37, no. 5, pp. 607-629. DOI: [org/10.1002/mas.21551](https://doi.org/10.1002/mas.21551).
8. Godzien J., De la Fuente A.G., Otero A., Barbas C. Metabolite Annotation and Identification. *Compr. Anal. Chem.*, 2018, vol. 82, pp. 415-445. DOI: [org/10.1016/bs.coac.2018.07.004](https://doi.org/10.1016/bs.coac.2018.07.004).
9. Molyneux R.J., Beck J.J., Colegate S.M., Edgar J., Gaffield W., Gilbert J., Hofmann T., McConnell L.L., Schieberle P. Guidelines for unequivocal structural identification of compounds with biological activity of significance in food chemistry (IUPAC Technical Report). *Pure Appl. Chem.*, 2019, vol. 91, no. 8, pp. 1417-1437. DOI: [org/10.1515/pac-2017-1204](https://doi.org/10.1515/pac-2017-1204).
10. Zubarev R.A., Makarov A. Orbitrap mass spectrometry. *Anal. Chem.*, 2013, vol. 85, no. 11, pp. 5288-5296. DOI: [org/10.1021/ac4001223](https://doi.org/10.1021/ac4001223).
11. Gruber B., Weggler B.A., Jaramillo R., Murrell K.A., Piotrowski P.K., Dorman F.L. Comprehensive two-dimensional gas chromatography in forensic science: A critical review of recent trends. *TrAC Trends Anal. Chem.*, 2018, vol. 105, pp. 292-301. DOI: [org/10.1016/j.trac.2018.05.017](https://doi.org/10.1016/j.trac.2018.05.017).
12. Nov kov  L., Svoboda P., Pavl k J. Ultra-high performance liquid chromatography. *Liquid Chromatography. Fundamentals and Instrumentation*. Elsevier, 2017, pp. 719-769. DOI: [org/10.1016/B978-0-12-805393-5.00029-4](https://doi.org/10.1016/B978-0-12-805393-5.00029-4).
13. Tang D.Q., Zou L., Yin X.X., Ong C.N. HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS. *Mass Spectrom. Rev.*, 2016, vol. 35, no. 5, pp. 574-600. DOI: [org/10.1002/mas.21445](https://doi.org/10.1002/mas.21445).
14. Bla enovi  I., Kind T., Ji J., Fiehn O. Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. *Metabolites*, 2018, vol. 8, no. 2, p. 31. Available at: <https://doi.org/10.3390/metabo8020031> (Accessed 21 May, 2020).
15. Matyushin D.D., Sholokhova A.Y., Karnaeva A.E., Buryak A. K. Various aspects of retention index usage for GC-MS library search: A statistical investigation using a diverse data set. *Chemom. Intell. Lab. Syst.*, 2020, vol. 202, p. 104042. Available at: <https://doi.org/10.1016/j.chemolab.2020.104042> (Accessed 21 May, 2020).
16. Milman B.L., Ostrovidova E.V., Zhurkovich I.K. Isomer differentiation using *in silico* MS<sup>2</sup> spectra. A case study for the CFM-ID mass spectrum predictor. *Mass Spectrom. Lett.*, 2019, vol. 10, no. 3, pp. 93-101. DOI: [10.5478/MSL.2019.10.3.93](https://doi.org/10.5478/MSL.2019.10.3.93).
17. Milman B.L., Zhurkovich I.K. The chemical space for non-target analysis. *TrAC Trends Anal. Chem.*, 2017, vol. 97, pp. 179-187. DOI: [org/10.1016/j.trac.2017.09.013](https://doi.org/10.1016/j.trac.2017.09.013).
18. Milman B.L., Zhurkovich I.K. Mass spectral libraries: A statistical review of the visible use. *TrAC Trends Anal. Chem.*, 2016, vol. 80, pp. 636-640. DOI: [org/10.1016/j.trac.2016.04.024](https://doi.org/10.1016/j.trac.2016.04.024).
19. *Analytical quality control and method validation procedures for pesticide residues analysis in food and feed (2019)*. Available at: [https://ec.europa.eu/food/sites/food/files/plant/docs/pesticides\\_mrl\\_guidelines\\_wrkdoc\\_2019-12682.pdf](https://ec.europa.eu/food/sites/food/files/plant/docs/pesticides_mrl_guidelines_wrkdoc_2019-12682.pdf) (accessed 21 May, 2020).
20. *Minimum criteria for chromatographic-mass spectrometric confirmation of the identity of analytes for doping control purposes (2015)*. Available at: [https://www.wada-ama.org/sites/default/files/resources/files/td2015idcr\\_-\\_eng.pdf](https://www.wada-ama.org/sites/default/files/resources/files/td2015idcr_-_eng.pdf) (accessed 21 May, 2020).
21. *Acceptance criteria for confirmation of identity of chemical residues using exact mass data for the FDA foods and veterinary medicine program (2015)*. Available at: <https://www.fda.gov/media/96499/download> (accessed 21 May, 2020).
22. *AORC guidelines for the minimum criteria for identification by chromatography and mass spectrometry (2016)*. Available at: <http://www.aorc-online.org/documents/aorc-ms-criteria-modified-23-aug-16>. (accessed 21 May, 2020).
23. *Guidance for industry: Mass spectrometry for confirmation of the identity of animal drug residues (2003)*. Available at: <https://www.fda.gov/media/70154/download>. (accessed 21 May, 2020).
24. Milman B.L., Zhurkovich I.K. Identification performance of low-molecular compounds by searching tandem mass spectral libraries with simple peak matching. *Mass Spectrom. Lett.*, 2018, vol. 9, no. 3, pp. 73-76. DOI: [10.5478/MSL.2018.9.3.73](https://doi.org/10.5478/MSL.2018.9.3.73).
25. *NIST 17 GC method / Retention index library*. Available at: <https://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:ridatabase> (accessed 21 May, 2020).
26. *CAS content*. Available at: <https://www.cas.org/about/cas-content> (accessed 21 May, 2020).
27. *PubChem*. Available at: <https://pubchem.ncbi.nlm.nih.gov> (accessed 21 May, 2020).
28. *ChemSpider*. Available at: <http://www.chemspider.com> (accessed 21 May, 2020).
29. *ZINC 15*. Available at: <http://zinc15.docking.org> (accessed 21 May, 2020).
30. Milman B.L., Ostrovidova E.V., Zhurkovich I.K. Bol'shie himicheskie bazy dannyh svobodnogo dostupa v netselevom mass-spektrometricheskom analize [Big chemical databases in a modern non-target mass spectrometry analysis]. *Mass-spektrometriia [Mass Spectrometry]*, 2020, vol. 17, no. 2, pp. 125-132 (In Russian).
31. *Wiley registry of mass spectral data*. Available at: <https://www.sisweb.com/software/wiley-registry.htm> (accessed 21 May, 2020).
32. *The NIST 20 Mass spectral library*. Available at: <https://www.sisweb.com/software/ms/nist.htm> (accessed 21 May, 2020).
33. Montenegro-Burke J.R., Guigas C., Siuzdak G. METLIN: a tandem mass spectral library of standards. *Computational Methods and Data Analysis for Metabolomics*. New York, Humana, 2020, pp. 149-163.
34. *METLIN*. Available at: [https://metlin.scripps.edu/ms\\_ms\\_spectrum\\_match\\_search.php](https://metlin.scripps.edu/ms_ms_spectrum_match_search.php) (accessed 21 May, 2020).
35. *GNPS: Global natural products social molecular networking*. Available at: <https://dorresteinlab.ucsd.edu/gnps> (accessed 21 May, 2020).
36. *MoNA - MassBank of North America*. Available at: <https://mona.fiehnlab.ucdavis.edu> (accessed 21 May, 2020).
37. *MassBank*. Available at: <http://www.massbank.jp> (accessed 21 May, 2020).